

A Light-weight Approach to Coreference Resolution for Named Entities in Text

Marin Dimitrov¹, Kalina Bontcheva², Hamish Cunningham², Diana Maynard²

¹Ontotext Lab, Sirma AI
38A Hristo Botev Blvd., Sofia 1000, Bulgaria
marin@sirma.bg

²Sheffield University
Regent Court, 211 Portobello Street, Sheffield S1 4DP, UK
kalina@dcs.shef.ac.uk, hamish@dcs.shef.ac.uk

Abstract

This paper presents a lightweight approach to pronoun resolution in the case when the antecedent is named entity. It falls under the category of the so-called "knowledge poor" approaches that do not rely extensively on linguistic and domain knowledge. We provide a practical implementation of this approach as a component of the General Architecture for Text Engineering (GATE). The results of the evaluation show that even such shallow and inexpensive approaches provide acceptable performance for resolving the pronoun anaphors of named entities in texts.

1. Introduction

Anaphora resolution and the more general problem of coreference resolution are very important for several fields of Natural Language Processing such as Information Extraction, Machine Translation, Text Summarization and Question Answering Systems.

Because of its importance, the problems are addressed in various works and many approaches exist (for an overview see e.g., (Mitkov, 1999)). The approaches differ in the approach they use (symbolic, neural networks, machine learning, etc.); the domain of the texts that they are tuned for; their comprehensiveness (e.g. is only pronominal anaphora considered); and in the results achieved.

This work falls under the class of "knowledge poor" approaches to pronominal anaphora resolution. Such methods are intended to provide inexpensive and fast implementations that do not rely on complex linguistic knowledge, yet they work with sufficient success rate for practical tasks (e.g., (Mitkov, 1998)).

Our approach is similar to other salience-based approaches, which perform resolution following the steps:

- identification of the antecedents in the context of the pronoun
- inspecting the context for candidate antecedents that satisfy a set of consistency restrictions
- assigning salience values to each antecedent based on a set of rules and factors
- choosing the candidate with the best salience value

The approaches that influenced our implementation were focused on anaphora resolution of certain set of pronouns in technical manuals. The goal of our work is resolution of pronoun anaphora in the case where the antecedent is a named entity - a person, organization, location, etc. The implementation relies only on the part-of-speech information, named entity recognition and orthographic coreferences existing between the named entities. No syntax parsing, focus identification or world-knowledge based approaches were employed. The texts that we used for the evaluation were newswire articles part

of the ACE (Automatic Content Extraction) competition training corpus (ACE, 2000). The evaluation showed that acceptable results could be achieved with such inexpensive approaches.

We provide an implementation of the approach, available as a component integrated with the General Architecture for Text Engineering (GATE) - a Language Engineering framework and set of tools developed by the University of Sheffield (Cunningham et al, 2002)¹.

2. Corpus Analysis

We used the ACE test corpora, which are of three different types, according to the source:

- *broadcast news programs* (BNEWS), generated with the help of automated speech recognition (ASR) systems. The news is from news programs of ABC News, CNN, VOA and PRI. Contains 60,000+ words.
- *newspaper* (NPAPER), generated by optical-character recognition (OCR) processing of newspaper sources. The corpus contains articles mainly from "The Washington Post". Contains 61,000+ words.
- *newswire* (NWIRE). Contains 66,000+ words.

We analysed these texts in order to have better understanding of the specifics related to each type of corpus. First we made an analysis of the pronoun distribution in the texts, and later an analysis of *pleonastic it* occurrences was performed. Not all pronouns were included in the analysis, only the following categories:

- *personal* - I, me, you, he, she, it, we, they, me, him, her, us, them.
- *possessive adjectives* - my, your, her, his, its, our, their.
- *possessive pronouns* - mine, yours, hers, his, its, ours, theirs.
- *reflexive* - myself, yourself, herself, himself, itself, ourselves, yourselves, themselves, oneself.

There were cases in which a pronoun can be classified in more than one category. For example "his" and "its"

¹ Available from <http://gate.ac.uk>.

may be possessive pronoun or possessive adjective. This is not a problem, since the part-of-speech (POS) tagger will identify this and will assign the proper category for the pronoun ("PRP" for possessive pronouns, and "PRPS" for possessive adjectives).

2.1. Total Pronouns

The percent of words that are pronouns reported in (Barbu & Mitkov, 2001) is 1.5% (422 pronouns out of 28,272 words). The average ratio we observed was almost three times higher. This is probably due to the specific differences in the domain of the analysed texts. The corpus in (Barbu & Mitkov, 2001) consists of technical manuals where specific grammatical constructs and language is being used. The ACE corpus consists of news articles and interviews where the number of named entities and the pronouns used to refer to them is unsurprisingly much higher.

The percentage of pronouns is shown in the following table:

source	Words	Pronouns	Pronouns (% of words)
npaper	61319	2264	3.7%
bnews	60316	3392	5.6%
nwire	66331	2253	3.4%
TOTAL	187966	7909	4.2%

Table 1. Number of pronouns and number of words in the ACE corpus

It is worth pointing out that the NWIRE and NPAPER part of the ACE corpus contain similar percentage of pronouns, while the percentage of pronouns in BNEWS is much higher. This is due to the fact that BNEWS contains mostly quotes speech dialogs, where pronouns are used more often than the names of the entities.

2.2. Distribution of pronouns by type

The relative distribution of pronouns by type is similar to the one reported in (Barbu & Mitkov, 2001). Again the most significant share is the one of the personal pronouns, followed by the possessive pronouns while the share of reflexive pronouns is insignificant (1.5%).

source	pronouns	pers.	pers. %	poss.	poss. %
npaper	2264	1593	70.4%	627	27.7%
bnews	3392	2862	84.4%	491	14.5%
nwire	2253	1629	72.3%	586	26.0%
TOTAL	7909	6084	76.9%	1704	21.5%

Table 2. Distribution of personal and possessive pronouns in the ACE corpus

The similarity between NPAPER and NWIRE corpora is observed again. The percentages for BNEWS are quite different from the rest and are closer to the ones reported in (Barbu & Mitkov, 2001).

The following table shows the relative importance of the 10 most often observed pronouns in each corpus:

npaper		bnews		nwire	
pronoun	%	pronoun	%	pronoun	%
HE	18.3%	IT	18.9%	IT	18.9%
IT	16.8%	I	11.6%	HE	16.5%
HIS	12.0%	YOU	11.6%	HIS	11.0%
ITS	8.6%	HE	10.5%	I	8.2%
THEY	8.0%	THEY	10.1%	THEY	8.1%
I	6.5%	WE	9.4%	ITS	6.7%
WE	6.4%	HIS	6.1%	WE	6.7%
SHE	4.8%	ITS	3.1%	YOU	5.0%
HER\$	3.3%	SHE	2.6%	SHE	2.6%
THEM	2.7%	HER\$	2.0%	HER\$	2.2%

Table 3. Relative importance of the 10 most often observed pronouns in different parts of the ACE corpus. HER\$ is the possessive adjective for SHE, not the object personal pronoun for SHE

There exists significant difference in the distribution of certain pronouns in different corpora. For example "I" and "you" and "we" which are expected to indicate quoted speech presence constitute around 13% and 19% of the pronouns in NPAPER and NWIRE respectively, while the percentage for BNEWS is almost twice as high - 32.6%.

Another fact of interest that is not shown in the table is the relative unimportance of possessive pronouns (*mine, yours, etc.*) in the text. There were only two such pronouns observed in the NPAPER corpus, constituting 0.1% of the pronouns, and there were no such pronouns in the BNEWS and NWIRE corpora. This implies that the coreference resolution algorithm may effectively ignore such pronouns because their (un)successful handling will not influence significantly the overall performance.

The same holds for reflexive pronouns. They constitute about 1.5% of the pronouns in the three corpora, so their effective resolution is unlikely to contribute sufficiently to good performance.

2.3. Pleonastic It Statistics

We analysed the three corpora for *pleonastic it* constructs. A full analysis for all non-anaphoric pronouns was out of the scope of this work. The percentage of *pleonastic it* occurrences we observed was low compared to the percentages reported by other researchers, e.g., 7.7% in (Lappin & Leass, 1994). This difference is most likely a consequence of the different domain of the analysed texts - technical manuals vs. news articles and interviews.

Note that the statistics for BNEWS and NWIRE are quite similar but they differ a lot from the ones for NPAPER. It is also worth pointing that *pleonastic it* constitutes a large percent of the total number of occurrences of "it" so if pleonastic pronouns are ignored in the implementation of the resolution algorithm, the final results for "it" are likely to be unsatisfactory. This is even more important if we consider that "it" constitutes about 19% of the pronouns in the three corpora.

	Pro-nouns	IT	pleon-It	pleon-It (% of pronouns)	pleon-It (% of IT)
npaper	2264	381	79	3.5%	20.7%
bnews	3392	642	105	3.1%	16.4%
nwire	2253	425	70	3.1%	16.5%
TOTAL	7909	1448	254	3.2%	17.5%

Table 4. *Pleonastic it* occurrences as nominal value, percentage of all pronouns, percentage of "it"

3. Design of the coreference resolution module

The analysis of the 3 ACE corpora helped us clarify and prioritise the requirements for the implementation of the module.

The coreference module has modular structure - it consists of a main module and a set of submodules. The main module takes care to initialise the submodules, to execute them in the specified order and finally to combine the results generated from the submodules and eventually to perform some post processing over the result.

This modular structure provides sufficient flexibility, so that the behaviour of the coreference module may be modified or tuned for specific tasks. Such specific tasks may require that the order in which submodules are executed may be changed (unless there are interdependencies between them). For certain tasks it may not be feasible to load and execute some modules at all if they are unlikely to contribute much for the final result. This is the case with technical manuals, which do not usually contain quoted speech fragments, so the submodule identifying such fragments in the text will not be useful.

The modular structure also makes it possible that new submodules be plugged in the main coreference modules when they become available. This is especially important for GATE because our intent is to extend the basic pronoun resolution functionality once certain lexical and ontological resources are integrated with the system (such integration is in progress at present).

Currently the main module consists of three submodules:

- *quoted text* module
- *pleonastic it* module
- pronoun resolution module

The *quoted text* submodule identifies quoted fragments in the text being analysed. The identified fragments are used by the pronoun resolution submodule for the resolution of pronouns such as *I*, *me*, *my*, etc. that appear in quoted speech fragments.

The submodule does not handle perfectly all the possible constructs of quoted fragments, which degrades the performance of the pronoun submodule. The main reason for this is the lack of correctly balanced quotation marks in the ACE corpora, especially the texts that were produced by OCR.

3.1. *Pleonastic It* submodule

The *pleonastic it* submodule is responsible for detecting pleonastic occurrences of "it".

As we already discussed above, the number of *pleonastic it* occurrences observed was significantly less than the numbers reported by other researchers. Yet the relative share of *pleonastic it*, as a percentage of all the occurrences of *it* makes identification of the former useful.

Previous work, such as (Lappin & Leass, 1994), contains patterns about *pleonastic it*. Unfortunately we discovered that these patterns would not be sufficient for all typical cases observed in our corpora:

- Often a synonym or antonym of a modal adjective or a synonym of a cognitive verb appears in the construct.
- The patterns are not flexible enough and miss even small variations of the defined constructs.
- It is unclear to which extent the patterns will deal with various syntactic variants of *be*.
- There are constructs in the ACE corpus which will not be matched by these patterns.

We resolved the first problem by adding synonyms and antonyms from WordNet to extend the set of modal adjectives and cognitive verbs from the basic set given in (Lappin & Leass, 1994).

The other problems had to be resolved by extending the base patterns – we used those in (Lappin & Leass, 1994):

1. **It be (adverb01) modaladj (conj01) S**
2. **It be (adverb01) modaladj (for NP) to VP**
3. **It is (adverb01) cogv-ed that S**
4. **It (adverb01) verb01 (conj02 | to) S**
5. **NP verb02 it (adverb01) modaladj (conj01 NP) to VP**

We dropped patterns 6 and 7 from the original paper, because they constituted less than one percent of the observed *pleonastic it* occurrences.

In the patterns above we have:

be = {be, become, remain}

adverb01 = {highly, very, still, increasingly, certainly, absolutely, especially, entirely, simply, particularly, quite, also, yet, even, more, most, often, rarely}

modaladj is the set of modal adjectives already discussed

conj01 = {for, that, is, whether, when}

conj02 = {that, if, as, like}

cogv-ed is the passive participle of the cognitive verbs defined above

verb01 = {seem, appear, look, mean, happen, sound}

verb02 = {find, make, consider}

Our implementation of these pattern extends the rules so that:

1. Different forms of the sets of verbs **be**, **verb01** and **verb02** are recognized (base, present 3rd person, present non-3rd person, past participle).
2. Question forms are matched.
3. Modal verbs used with the above sets are matched.
4. Negation is matched.

We identified one more pattern that was observed often in the ACE corpus, but we did not implement it, because

the pattern was not generic enough and depends on too many specific expressions. The pattern looks like

6. It be/take time-expr before/since S

...where **time-expr** represents time expressions such as *two weeks, today, one month, a while, longer*, etc.

The following table lists the distribution of the pleonasm from each type observed in the ACE corpora together with the percentage of the occurrences correctly identified.

Pattern	Occurrences	% of pleonastic it	identified
1	35	13.9%	72.0%
2	65	25.8%	72.0%
3	3	1.2%	33.3%
4	18	7.1%	77.8%
5	11	4.4%	72.7%
6	16	6.3%	-
Unclass.	104	41.3%	-
TOTAL	252		37.7%

Table 5. Pleonastic-it statistics

Note that patterns 1 and 2 are observed most often and the percentage of pleonastic it constructs that were not matched by any pattern is very high - more than 40%. The precision (number of occurrences matched / all occurrences of this type) of the specific rules is relatively good and with the exception of one rule it is more than 70% but the high number of unclassified occurrences degrade the overall performance.

3.2. Pronoun Resolution Submodule

The main functionality of the coreference resolution module is in the pronoun resolution submodule. This submodule uses the result from the execution of the quoted speech and *pleonastic it* submodules.

The module works according to the following algorithm:

1. For each pronoun:
 - inspect the appropriate context for all candidate antecedents for this kind of pronoun;
 - choose the best antecedent (if any).
2. Create the coreference chains from the individual anaphor/antecedent (this step is performed from the main coreference module).

Pronoun resolution (step 1) works as follows:

- If the pronoun is *it* then a check is performed if this is a pleonastic occurrence and if so then no further attempt for resolution is made.
- The proper context is determined. The context size is expressed in the number of sentences it will contain. The context always includes the current sentence (the one containing the pronoun), the preceding sentence and zero or more preceding sentences.
- Depending on the type of pronoun a set of candidate antecedents is proposed. The candidate set includes the named entities that are compatible

with this pronoun. For example if the current pronoun is *she* then only the Person entities with gender "female" or "unknown" will be considered as candidates. From all candidates one is chosen according to evaluation criteria specific for the pronoun.

3.2.1. Resolution of *she, her, her\$, he, him, his, herself, himself*

The resolution of *she, her, her\$, he, him, his, herself* and *himself* is similar because the analysis of the corpus showed that these pronouns are related to their antecedents in similar manner. The characteristics of the resolution process are:

- Context inspected is not very big - cases where the antecedent is found more than 3 sentences further back than the anaphor are rare.
- Recency factor is heavily used - the candidate antecedents that appear closer to the anaphor in the text are scored better.
- Anaphora has higher priority than cataphora. If there is an anaphoric candidate and a cataphoric one then the anaphoric one is preferred, even if the recency factor scores the cataphoric candidate better.

The resolution process performs the following steps:

1. Inspect the context of the anaphor for candidate antecedents. Each Person entity is considered as a candidate. Cases where *she/her* refers to inanimate entity (ship for example) are not handled.
2. For each candidate, perform a gender compatibility check - only candidates having "gender" feature equal to "unknown" or gender compatible with the pronoun are considered for further evaluation.
3. Evaluate each candidate against the best candidate so far:
 - If the two candidates are anaphoric for the pronoun then choose the one that appears closer.
 - The same holds for the case where the two candidates are cataphoric relative to the pronoun.
 - If one is anaphoric and the other is cataphoric then choose the former, even if the latter appears closer to the pronoun.

3.2.2. Resolution of *it, its, itself*

This set of pronouns also shares many common characteristics. The resolution process contains certain differences with the one for the previous set of pronouns.

Successful resolution for *it, its, itself* is more difficult because of the following factors:

- There is no gender compatibility restriction. In the case when there are several candidates in the context, the gender compatibility restriction is very useful for rejecting some of the candidates. When no such restriction exists, and with the lack of any syntactic or ontological information about the

² This is the possessive adjective for *she*, not the object for of the personal pronoun.

entities in the context, the recency factor plays the major role for choosing the best antecedent.

- The number of nominal antecedents (i.e. entities that are referred not by name) is much higher compared to the number of such antecedents for *she*, *he*, etc. In this case trying to find antecedent only amongst named entities degrades the precision substantially.

We performed analysis of the occurrences of *it*, *its* and *itself* in the ACE corpus in order to determine the usefulness of the recency factor if it is used as the only factor for choosing the best antecedent:

- In 52% of the cases the most recent named entity of type Organization and Location was the correct antecedent.
- In 15% of the cases the antecedent was a named entity which was not the most recent related to the anaphor.
- In 33% of the cases the antecedent was nominal where the resolution will fail, as their resolution is not yet supported.

The analysis shows that the recency factor all by itself offers some means of correct pronominal resolution. Further, we identified that half of the cases (7.3%) where the antecedent was not the most recent named entity were appositional. For example:

Yamaichi Securities Co., once *Japan*₂'s largest securities house, officially closed *its*₁ last offices today after authorities revealed the severity of its losses.

In this example if the best antecedent is chosen on the basis of recency then *its* will be incorrectly matched to *Japan*. If apposition was identified, then the most proper choice would have been the named entity to which the apposition itself refers (in this case *Yamaichi Securities Co.*).

The resolution steps are similar to the ones for *she*, *he*, etc. with the following differences:

- Entities of type Location and Organizations are considered as candidate antecedents.
- Only recency is considered for choosing the best antecedent.
- Named entities that are cataphoric to the pronoun are not considered as candidate antecedents.

3.2.3. Resolution of *I*, *me*, *my*, *myself*

Resolution of these pronouns is dependent on the work of the quoted speech submodule. One important difference from the resolution process of other pronouns is that the context is not measured in sentences but depends solely on the quote span. Another difference is that the context is not contiguous - the quoted fragment itself is excluded from the context, because it is unlikely that an antecedent for *I*, *me*, etc. appears there. The context itself consists of:

- The part of the sentence where the quoted fragment originates that is not contained in the quote - i.e. the text prior to the quote.
- The part of the sentence where the quoted fragment ends that is not contained in the quote - i.e. the text following the quote.

- The part of the sentence preceding the sentence where the quote originates, which is not included in another quote.

For example the context for the following example is underlined:

Others believe things will move more slowly. "I don't expect to see a significant change as of April 1," said Mitsuru Saito, market economist for Sanwa Bank.

Another difference from the pronouns in the first group (*he*, *she*, etc.) is that candidate antecedents are considered to be not only entities of type Person but also the pronouns *he* and *she*.

We identified several patterns that classify the relation between the pronouns *I*, *me*, *my*, *myself* and their antecedents. The subset of the corpus that was analysed consisted of 40 documents containing 95 quoted fragments with 72 occurrences of the pronouns of interest. The patterns we identified for these 72 occurrences are:

- The antecedent is the closest named entity in the text following the quoted fragment. This pattern is observed in 52% of the cases. An example is: "*I* did not urge anyone to say anything that was untrue," *Clinton*₁ told Lehrer.
- The antecedent is found in the sentence preceding the sentence where the quoted fragment originates. If the preceding sentence also contains a quote then the antecedent is usually the named entity (or pronoun) that is the one most close to the end of the quote. This pattern was observed in 29% of the cases. An example is: "*I* did not urge anyone to say anything that was untrue," *Clinton*₁ told Lehrer. "That's *my*₁ statement to you"
- The antecedent is the closest named entity preceding the quote in the sentence where the quote originates. This pattern counts for less than 3% of the cases. An example is: *U.S. officials said there was confusion about whether China would fulfill the contracts, but Cohen*₁ declared: "*I* believe we have assurances that such sales will not continue."
- The antecedent is either nominal (13%) or a named entity in position where the patterns above will not identify it correctly (3%). These cases will not be handled correctly by the algorithm.

It is worth noting that contrary to other pronouns, the antecedent for *I*, *me*, *my* and *myself* is most often cataphoric or if anaphoric it is not in the same sentence with the quoted fragment.

The resolution algorithm consists of the following steps:

1. Locate the quoted fragment description that contains the pronoun. If the pronoun is not contained in any fragment then return without proposing an antecedent.
2. Inspect the context of the quoted fragment (as defined above) for candidate antecedents.
3. Try to locate a candidate in the text succeeding the quoted fragment (first pattern). If more than one candidate is present, choose the closest to the end

of the quote. If a candidate is found then choose it as an antecedent.

4. Try to locate candidate in the text preceding the quoted fragment (third pattern). Choose the closest one to the beginning of the quote. If found then choose as an antecedent.
5. Try to locate antecedents in the unquoted part of the sentence preceding the sentence where the quote starts (second pattern). Give preference to the one closest to the end of the quote (if any) in the preceding sentence or closest to the sentence beginning.

4. Evaluation

We manually annotated a subset of the ACE corpora in order to evaluate precision, recall and F-measure for the implementation. The subset consists of 21 randomly selected documents (7 from each corpus) containing 352 pronouns. The evaluation corpus represents 5% of the documents in the ACE corpus and contains 4.5% of the pronouns. No pronouns were excluded from the evaluation. Occurrences of pronouns that are not handled yet degrade the recall. Nominal antecedents degrade the precision. The results were 66% precision and 46% recall. These numbers are comparable to the performance of other knowledge-poor coreference resolution approaches, e.g., (Barbu & Mitkov, 2001; Mitkov, 1998). The following table contains the results for each individual group of pronouns:

Pronoun group	precision	recall	f-measure
1	79.3%	77.2%	78.2%
2	43.5%	51.7%	47.2%
3	77.8%	62.2%	70%

Table 4. Precision, recall and f-measure for the three groups of pronouns (1st group includes *he, she*, etc., the 2nd group includes *it, its* and *itself*, the 3rd one includes *I, me, myself* and *my*)

The results show that the resolution of pronouns such as *he, she, her*, etc. is relatively successful even with such simple heuristic patterns used and without incorporating any syntax or centering information. The precision is degraded by the nominal antecedents. The algorithm will also benefit from some syntax information indicating the subject of the sentence, because the results show that the recency factor and the gender agreement are not sufficient.

The resolution of pronouns such as *it, itself* and *its* is less successful. Apart from nominal antecedents which have even greater impact for this group, additional degradation is induced from the low performance of the *pleonastic it* module, which, although using rules that cover more cases than the ones in (Lappin & Leass, 1994), still identifies only 38% of the pleonastic occurrences. It is worth noting that the *pleonastic it* module has very high precision but low recall, so further extension of its patterns will improve the recall and will have positive impact on the resolution of *it*.

The recall errors for resolution of *I, me*, etc. are mainly due to errors made by the quoted speech submodule. Additionally the performance is negatively impacted by the specifics of the BNEWS corpus, where the quoted fragments are not marked in the text, and as a result no attempt for resolution of the pronouns of the 3rd group will be made on this corpus.

Finally, if we measure the performance of the coreference module, independently from the named entity recogniser, i.e., against the same corpus but with manually annotated named entities, then precision goes up to 73% and recall up to 53% for all pronouns, with the biggest improvement for group 3 where precision goes up to 86% and recall is 76%.

5. Conclusion

The lightweight approach presented here achieves acceptable performance without using any syntax information or centering theory methods, which shows that even simple heuristic rules identified from corpus analysis can be sufficient for simple coreference functionality.

Unfortunately further improvement in precision and recall by incorporating lightweight techniques is unlikely to be achieved. That is why we intend to incrementally extend the basic functionality with new features. In future work, we will address apposition identification, extending the set of handled pronouns, and a module for resolving nominal anaphora.

6. References

- ACE (2000). *Entity Detection and Tracking - Phase 1*. ACE Pilot study definition available from the ACE site <http://www.itl.nist.gov/iaui/894.01/tests/ace/index.htm>.
- Barbu, C. & Mitkov, R. (2001). Evaluation tool for rule-based anaphora resolution methods. *Proceedings of ACL'01*, Toulouse, 2001.
- Cunningham, H. & Maynard, D. & Bontcheva, K. & Tablan, V. (2002). GATE: A framework and graphical development environment for robust NLP tools and applications. *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'2002)*.
- Lappin, S. & Leass, H. (1994). An algorithm for pronominal anaphora resolution, *Computational Linguistics*, 20(4).
- Mitkov, R. (1998). Robust Anaphora Resolution with Limited Knowledge. *Proceedings of COLING'98 /ACL'98*.
- Mitkov, R. (1999). Anaphora Resolution: The State of the Art. Working paper, University of Wolverhampton.